

МАКЕДОНСКИОТ КОРПУС НА АНГЛИСКИОТ МЕЃУЈАЗИК И НЕГОВАТА ПРИМЕНА ВО ЛИНГВИСТИКАТА И ВО ЛИНГВОДИДАКТИКАТА

АПСТРАКТ: Во оваа статија го претставуваме Македонскиот корпус на англискиот меѓујазик (МКАМ). Тој претставува електронска база на текстови пишувани од изучувачи на англискиот јазик во Република Македонија. Создаден е во рамките на проектот *Анализа на англискиот меѓујазик од македонски говорители* на А1, А2 и Б1 ниво од Заедничката европска референтна рамка за јазиците (ЗЕРР) со посебен осврт на меѓујазичното влијание. Целите на проектот се да се опишат карактеристиките на англискиот меѓујазик од македонски говорители. Електронската база е создадена како основа за емпириските истражувања.

Во статијата се опишува врз кои принципи се базира електронскиот корпус и какви можности нуди за истражувачите и за наставниците по англиски јазик. Притоа ќе бидат прикажани и некои методи за истражување на меѓујазикот користени во досегашните студии врз материјалот од корпусот. На крајот се разгледуваат можностите за понатамошно дополнување, развивање и усовршување на корпусот.

Клучни зборови: електронски корпус, меѓујазик, методи на истражување, меѓујазично влијание, анализа на грешки, контрастивна анализа

1. Вовед

Македонскиот корпус на англискиот меѓујазик (МКАМ) е електронска база на текстови пишувани од изучувачи на англискиот јазик во РМ. Базата е создадена во рамките на проектот „Анализа на англискиот меѓујазик од македонски говорители на А1, А2 и Б1 ниво од Заедничката европска референтна рамка за јазиците (ЗЕРР) со посебен осврт на меѓујазичното влијание“, кој беше прифатен од МОН (Проект Бр. 13-3965/1) за 2010-2012 година. Во текот на работата опсегот на истражувањата се прошири и на Б2 ниво. Во проектот учествува истражувачи кои предаваат на повеќе универзитети во Р.М.¹ и студенти на прв и втор циклус на студии. Општата цел на проектот е да се идентификуваат, класифицираат и објаснат некои карактеристики на англискиот меѓујазик од македонски говорители, поточно англискиот јазик кој го зборуваат изучувачите во Република Македонија. Поконкретно, истражувањата се насочени кон

¹ Главен истражувач: проф. д-р Лилјана Митковска, Универзитет ФОН, Скопје; Соработници – истражувачи: проф. д-р Елени Бужаровска, Филолошки факултет „Блаже Конески“, Скопје, доц. д-р Марија Кусевска, Универзитет „Гоце Делчев“, Штип и ас. м-р Анастасија Киркова-Наскова, Филолошки факултет „Блаже Конески“, Скопје; Информатичар: доц. д-р Славчо Чунгурски, Универзитет ФОН, Скопје.

определени структури на различни нивоа (морфо-синтаксичко, лексичко, прагматичко) со цел да се определи:

- (1) јазичната компетенција на поединечните нивоа: до кој степен е испитуваната структура усвоена на нивоата A1, A2, B1 и B2;
- (2) темпото на развој на дадената структура од едно ниво во друго: како напредуваат учениците во однос на усвојување на структурата со преминот од едно ниво во друго;
- (3) кои се проблемите на секое ниво: кои од нив се резултат на влијание на внатрешнојазични фактори (развојни грешки), а кои на влијание од J1 (преносни грешки).

При објаснувањето се обрнува повеќе внимание на преносните грешки. Испитувањето се фокусира на типичните проблеми кои ги имаат македонските изучувачи на англискиот јазик со цел за нивно поефикасно надминување. Со тоа овие истражувања се ставаат во функција на педагошката пракса. Резултатите од истражувањата треба да ги упатат наставниците на специфичните проблеми на македонските говорители во изучувањето на англискиот јазик, за кои не добиваат информација во општата литература за усвојување на англискиот јазик. Тоа треба да им помогне во планирањето на наставата со педагошкиот материјал што сега се користи, а е наменет за говорители со најразлични J1 и не води сметка за специфичните потреби на конкретни говорни групи.

2. Електронскиот корпус на јазикот на изучувачите

МКАМ беше создаден како дел од проектот за поефикасно спроведување на истражувањата на јазикот на изучувачите врз автентичен материјал. Покрај неговата функција во проектот, создавањето на МКАМ беше мотивирано и од следните подолгорочни цели:

- (1) Создавање на електронски корпус на англискиот меѓујазик којшто во иднина ќе може да се користи за други истражувања од областа на изучувањето на јазикот.
- (2) Развој на методологија на создавање на таков корпус, којшто ќе има поширока примена и за корпуси со други цели и за други јазици.
- (3) Развивање методологија на емпириски истражувања на јазикот на изучувачите на странски/втор јазик за практични цели.
- (4) Заживување на научно-истражувачката пракса со современи истражувачки методи и информатички технологии и оспособување на млад кадар за нејзино продолжување. Запознавање со најновите методи за јазични истражувања базирани на електронски корпуси и дефинирање на теми за истражувања кои треба/може да се работат врз основа на создадениот корпус.

(5) Поврзување на науката со праксата.

Електронските корпуси на јазикот на изучувачите започнуваат во 1980-те, но стануваат попопуларни последниве 10 години. Прв таков корпус е *Longman Learner Corpus*, кој се состои од 1 мил зборови и не е широко достапен. Тој се користи за целите на составувачите на речници, граматики, наставни материјали за англиски јазик и тестови. Сега постојат многу различни корпуси на јазикот на изучувачите. Во базата на податоци за вакви корпуси на Центарот за англиска корпус лингвистика (The Centre for English Corpus Linguistics - CECL) во Лувен, Белгија, *Learner corpora around the world* (<http://www.uclouvain.be/en-cecl-lcworld.html>), се наведени 121 корпус од различен тип: пишувани, говорни, за определено ниво, за повеќе нивоа итн. Повеќето се на изучувачи на англискиот јазик со различни L1 (француски, германски, јапонски, кинески, малезиски, чешки, руски и др.; со повеќе јазици), но има и за француски, германски и за други јазици. Корпусот создаден во рамките на овој центар, *International Corpus of Learner English*,² започнат во 1989 изврши големо влијание во создавањето методологија за креирање и аотирање на вакви корпуси, како и за нивна практична употреба во истражувањата и во педагошката пракса.

Што претставува „електронски корпус на јазик на изучувачи“? Секој збир на текстови од изучувачи не може да се сметан за таква база. Според GRANGER (2002: 7)

„Компјутерски корпуси на јазик на изучувачи се електронски бази на автентични текстови на странски/втор јазик собрани според експлицитни критериуми за определени цели. Тие се кодирани на стандардизиран и изедначен начин и нивниот извор и потекло се документирани.“³

Тоа значи дека, за да може ефикасно да се користи за различни истражувања, електронскиот корпус треба добро да се осмисли од аспект на податоците за испитаниците и природата на податоците што се собираат. Имајќи го тоа предвид, МКАМ го конципиравме врз определени принципи и јасно дефинирани параметри кои овозможуваат негово користење за истражувачки и други цели.

2.1 Собирање на податоци

Во периодот октомври - декември 2010 и 2011 година тимот работеше на првата фаза:

(1) конципирање и изработка на електронска база на податоци, и (2) собирање и членовите на тимот и студенти од повеќе градови на Македонија запишани на средни

² Прва верзија од корпусот е издадена електронски во 2002, а втората верзија, која содржи и аотирање на грешки, во 2009 (GRANGER et al. 2009)

³ „Computer learner corpora are electronic collections of authentic FL/SL textual data assembled according to explicit design criteria for a particular SLA/FLT purpose. They are encoded in a standardised and homogeneous way and documented as to their origin and provenance.“ (GRANGER, 2002: 7)